

REMARKS

This reply encompasses a bona fide attempt to overcome the rejections raised by the Examiner and presents amendments as well as reasons why the applicants believe that the claimed invention is novel and unobvious over the closest prior art of record, thereby placing
5 the present application in a condition for allowance.

Regarding Claim Status

Claims 1-124 were presented for examination. Claims 1-124 were objected and rejected. Claims 1-3, 7, 12-13, 21-22, 29-30, 35-39, 42-43, 45-46, 48-49, 52-56, 60-64, 92, 103-104,
10 and 116 are amended herein. Support for the amendments presented herein can be found in the specification as originally filed, *see, e.g.*, page 8, line 11, through page 9, line 17; page 12, lines 13-20; and page 31, lines 4-9. Claims 93-102 are cancelled, rendering the rejections with respect to these claims moot. By this Amendment, claims 1-92 and 103-124 are pending.

15

Regarding Claim Objections

Claims 1-124 were objected to for reciting “fingerprints” therein. More specifically, “fingerprints” were deemed “not a correct term of art for features in an audio sample.
20 ‘Voiceprints’ should be used.” Applicants respectfully disagree. According to the American Heritage® Dictionary, a “voiceprint” is defined as an electronically recorded graphic representation of a person's voice, uniquely characteristic of the individual speaker. Since the

present invention is applicable to all kinds of sounds and media, the use of “voiceprints” would be incorrect and misleading.

Patent Examiners must rely on the applicant's disclosure to properly determine the meaning of terms used in the claims. *Markman v. Westview Instruments*, 52 F.3d 967, 980, 34 USPQ2d 1321, 1330 (Fed. Cir.) (en banc), *aff'd*, U.S., 116 S. Ct. 1384 (1996).

An applicant is entitled to be his or her own lexicographer, and in many instances will provide an explicit definition for certain terms used in the claims. Where an explicit definition is provided by the applicant for a term, that definition will control interpretation of the term as it is used in the claim. *Toro Co. v. White Consolidated Industries Inc.*, 199 F.3d 1295, 1301, 53 USPQ2d 1065, 1069 (Fed. Cir. 1999) (meaning of words used in a claim is not construed in a “lexicographic vacuum, but in the context of the specification and drawings.”).

The recitation of “fingerprints” in the claims is consistent with the teaching of the application disclosure. The present application explicitly discloses that a “fingerprint” characterizes one or more features of a media sample or file at or near a particular location called “landmark” [Spec. page 8, lines 20-30]. Applicable types of media include, but not limited to, text, audio, video, image, and any multimedia combinations thereof [Spec. page 7, lines 24-26]. Thus, fingerprints in embodiments of the present invention are word strings, spectral components, or pixel RGB values [Spec. page 8, lines 28-30]. They are not “voiceprints” per se. Withdrawal of the claim objections is therefore respectfully requested.

Regarding 35 U.S.C. § 102 Rejections

Claims 107-111 were rejected under 35 U.S.C. § 102(b) as being anticipated by Gill *et al.* (U.S. Patent No. 4,415,767, hereinafter referred to as “Gill”). The rejections are respectfully traversed. Reconsideration is earnestly requested.

5

“A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.” *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987).

10

“The identical invention must be shown in as complete detail as is contained in the ... claim.” *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989).

15 Independent claim 107 is recited below for the convenience of the Examiner:

A method of characterizing an audio sample, comprising
 computing at least one fingerprint from a spectrogram of said audio
sample, wherein
 said spectrogram comprises an anchor salient point and linked salient
points, and wherein
 said fingerprint is computed from frequency coordinates of said anchor
salient point and at least one linked salient point.

20

Gill is hereby distinguished at least because Gill does not teach, neither expressly nor
25 inherently, “fingerprint” as set forth in claim 107. Moreover, Gill does not show or describe
an identical invention as is contained in the claims.

As a whole, Gill teaches a method and apparatus for *speech* recognition and reproduction. More specifically, Gill expressly teaches that a “voiceprint is defined by a sequence of acoustic features” [col. 17, lines 3-4, emphasis added]. This “voiceprint” is not a “fingerprint” that characterizes one or more features at or near a distinctive and reproducible landmark, as taught and claimed in the present patent application.

According to Gill, these “acoustic features” are time-averaged spectral amplitudes and time-rates-of-change of spectral amplitudes [col. 17, lines 4-6]. In other words, a voiceprint is defined by a sequence of time-averaged spectral amplitudes and time-rates-of-change of spectral amplitudes. It is not a fingerprint computed from frequency coordinates of linked salient points that are local maxima, local minima, zero crossings, or other distinctive features [Spec. page 16, line 32, through page 17, line 11].

Having distinguished Gill, claim 107 is submitted to be patentable under 35 U.S.C. § 102(b). Reliance is placed on *In re Fine*, 5 USPQ 2d 1596, 1600 (Fed. Cir. 1988) and *Ex parte Kochan*, 131 USPQ 204 (Bd. App. 1960) for the allowance of dependent claims 108-115 since they differ in scope from their parent independent claim 107 which is submitted to be patentable.

Claims 1-17, 35-37, 50-52, 54-55, 60-80, 92-93, 103-106, and 116-120 were rejected under 35 U.S.C. § 102(e) as being anticipated by Kanevsky *et al.* (U.S. Patent No. 6,434,520, hereinafter referred to as “Kanevsky”). The rejections are respectfully traversed. Reconsideration is earnestly requested in view of the amendments presented herein and the following remarks. Traversal to the rejections will be collectively discussed below with respect to independent claims 1, 55, 62-64, 103-104, and 116, which recite:

1. Computing/obtaining sample/file fingerprints;
2. Each fingerprint characterizes one or more features at or near a particular location (landmark) within a media sample/file;
3. Generating correspondences of values exceeding a threshold between sample landmarks and file landmarks, wherein the corresponding landmarks have equivalent (identical or similar) fingerprints; and
4. Identifying a winning media file that has a plurality of correspondences that are substantially linearly related.

Kanevsky is hereby distinguished at least because Kanevsky does not teach, neither expressly nor inherently, any of the above elements, including “fingerprints”, as set forth in the claims. Moreover, Kanevsky does not show or describe an identical invention as is contained in the claims.

As a whole, Kanevsky is directed to a system and method for indexing and querying audio archives. Using known methods, a feature extraction module 102 processes a stream of audio data on a frame-by-frame basis to generate a plurality of feature vectors [col. 3, lines 12-15].

Using conventional methods, a segmentation module 103 processes these feature vectors by determining and time-stamping the locations in the stream of feature vectors where changes in speaker, channel and/or background occur [col. 3, lines 22-26].

5 As such, the audio data stream is divided into a plurality of segments 104, each representing time intervals of the audio data stream having speech of distinct speakers, music, different channels and/or different background conditions [col. 4, lines 54-58]. A speaker identification/verification module 105 then uses pre-enrolled speaker models (or voiceprints) to perform speaker identification for each and every segment 104 of the entire audio data
10 stream [col. 4, line 66, through col. 5, line 5].

Once the segmentation and characterization of each segment has been performed, there is no further interaction between the segments. In Kanevsky, each segment is treated as an independent object in each file of the database and is thus independently considered for a
15 match according to the Boolean query attributes.

In contrast, the present invention as claimed in these claims does not perform any segmentation that divides an audio data stream into a plurality of segments based on feature vectors extracted frame-by-frame from the audio data stream. One may do so if desired in a
20 particular embodiment, but that is not what is called for in these claims. A key novelty of the present invention as claimed in these claims is the technique of testing and determining whether the corresponding sample and file “particular locations” (also referred to as

“landmarks”) have a linear relationship [*see*, Spec. page 23, line 29; and FIG. 10A]. Nothing in the applied art addresses this approach.

As a whole, the claimed invention is directed to a robust method of recognizing a media sample even when the signals to be recognized are subject to linear and non-linear distortions. This is possible in part because it is not the features or characteristics of the sample that are deterministic to the identifying step, but the *correspondences* between locations within the sample (i.e., sample landmarks) and locations within the recorded files (i.e., file landmarks).

More specifically, embodiments of the invention seek to identify a linear correspondence of the form $(\text{Landmark}^*_n = m \cdot \text{landmark}_n + \text{offset})$ between certain sample locations and certain file locations [*id.*]. Each of these “locations”, also referred to as “landmarks,” is associated with a “fingerprint” and each fingerprint characterizes one or more *features* of a media sample or media file at or near a landmark thereof. Candidate media files have file landmarks whose associated fingerprints are identical or similar (i.e., equivalent) to the sample fingerprints of the media sample.

The presence of a significant linearly aligned set of corresponding sample and file landmarks (e.g., corresponding timepoints as shown in FIG. 10A) yields the unexpected result that short audio samples (approximately 5-10 seconds) with heavy distortion and noise can be found/identified very quickly in an extremely large database of audio recordings (e.g., over 750,000,000 seconds, in over 2,500,000 unique recordings). Moreover, since the linearly

aligned set of corresponding sample and file landmarks are determined from statistically significant accumulation of differences in a histogram, they are not affected by and thus are immune to distortions, such as additive noise, and nonlinearities, such as voice compression artifacts. This is a major reason for the robustness of typical embodiments of the invention.

5 Nothing in the applied art teaches this. Further, at the time of the invention, such a robust and fast audio recognition technique was not in the general knowledge of one of ordinary skill in the art.

Not only Kanevsky does not teach or disclose identifying audio files, Kanevsky also does not
10 teach or disclose searching and retrieving audio files based upon arbitrary excerpts (audio samples). Instead, Kanevsky deals with indexing and retrieving well-defined independent segments of audio files. It seems highly unlikely, if not impossible, that Kanevsky's invention can correctly and very quickly identify audio files in an extremely large database of audio recordings from short, arbitrary audio samples.

15

Accordingly, it is respectfully submitted that the present invention as set forth in the independent claims 1, 55, 62-64, 103-104, and 116 recites subject matter not reached by the closest prior art of record, Kanevsky, under 35 U.S.C. §§ 102(e) and 103(a) and therefore should be allowed. Reliance is placed on *In re Fine*, 5 USPQ 2d 1596, 1600 (Fed. Cir. 1988)
20 and *Ex parte Kochan*, 131 USPQ 204 (Bd. App. 1960) for the allowance of dependent claims 2-54, 56-61, 65-91, 105-106, and 117-124 since they differ in scope from their respective parent independent claims which are submitted to be patentable.

Regarding 35 U.S.C. § 103 Rejections

Claims 18-34, 38-49, 53, 56-59, 81-91, and 121-124 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Kanevsky. The rejections are respectfully traversed.

5

When applying 35 USC 103, the following tenets of patent law must be adhered to:

- (A) The claimed invention must be considered as a whole;
- (B) The references must be considered as a whole and must suggest the desirability and thus the obviousness of making the combination;
- 10 (C) The references must be viewed without the benefit of impermissible hindsight vision afforded by the claimed invention; and
- (D) Reasonable expectation of success is the standard with which obviousness is determined.

15

Hodosh v. Block Drug Co., Inc., 786 F.2d 1136, 1143 n.5, 229 USPQ 182, 187 n.5 (Fed. Cir. 1986). *See also*, MPEP 2141.

20

The claimed invention is submitted to be patentable over Kanevsky at least according to (B). As the Examiner correctly pointed out on pages 6 and 7 of the Office Action, Kanevsky does not show or describe many claimed elements. Nor does Kanevsky suggest the desirability of making a combination that includes these claimed elements.

It is noted that the Examiner took Official Notices on claimed elements that Kanevsky neither teaches nor suggests. Specifically, the Examiner contended that, at the time the invention was made, one of ordinary skill in the art would have known that

1. various features (fingerprints) could be computed from frequency information, that Kanevsky could be modified as such, and that the resulting modification would have somehow arrived at an invention as set forth in claims 18-34, 81-91, and 121-124;
2. subsets of additional media files having various probabilities can be ranked, that Kanevsky could be modified as such, and that the resulting modification would have somehow arrived at an invention as set forth in claims 38-49; and
3. rolling buffers could be added to any processing system depending on the users preferences, that Kanevsky could be modified as such, and that the resulting modification would have somehow arrived at an invention as set forth in claims 53 and 56-59.

These Official Notices were not supported by documentary evidence [MPEP 2144.03]. It is respectfully submitted that there is no concrete evidence in the record that supports these findings. Moreover, the facts asserted to be well-known are submitted to be not capable of instant and unquestionable demonstration as being well-known.

If such notice is taken, the basis for such reasoning must be set forth explicitly. The examiner must provide specific factual findings predicated on sound technical and scientific reasoning to support his or her conclusion of common knowledge. See *Soli*, 317 F.2d at 946, 37 USPQ at 801; *Chevenard*, 139 F.2d at 713, 60 USPQ at 241.

If the examiner is relying on personal knowledge to support the finding of what is known in the art, the examiner must provide an affidavit or declaration setting forth specific factual statements and explanation to support the finding. See 37 CFR 1.104(d)(2).

Applicants specifically traverse the Official Notices and respectfully request that the Examiner, if the grounds of rejections are to be maintained, provides an affidavit or declaration setting forth specific factual statements and explanation to support the finding of
5 facts asserted to be well-known at the time of the invention.

Kanvesky is a fundamentally different invention. Kanvesky cannot and does not solve the audio recognition problem addressed by embodiments of the present invention. Kanvesky's "voiceprints" do not anticipate or suggest "fingerprints" as taught and claimed in the present
10 application.

To establish *prima facie* obviousness of a claimed invention, all the claim limitations must be taught or suggested by the prior art. *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). "All words in a claim must be considered
15 in judging the patentability of that claim against the prior art." *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970).

Accordingly, claims 18-34, 38-49, 53, 56-59, 81-91, and 121-124 are submitted to be patentable over Kanvesky under 35 U.S.C. § 103(a).

20

Conclusion

For the foregoing reasons, it is respectfully submitted that the claimed invention recites subject matter not reached by Kanvesky and Gill, under 35 U.S.C. §§ 102(b), 102(e), and 103(a), respectively. Favorable consideration and a Notice of Allowance of all pending
5 claims 1-92 and 103-124 are therefore earnestly solicited.

This Response/Amendment is submitted to be complete and proper in that it places the present application in a condition for allowance without adding new matters. The Examiner's attention is respectfully directed to the information disclosure statement (IDS) accompanying
10 this Response. Please consider all of the references cited thereon. Thank you.

The Examiner is sincerely invited to telephone the undersigned at 650-331-8413 for discussing an Examiner's Amendment or any suggested actions for accelerating prosecution and moving the present application to allowance.

15

Respectfully submitted,



Katharina Wang Schuster, Reg. No. 50,000
Attorney for the Applicants

LUMEN INTELLECTUAL PROPERTY SERVICES
2345 Yale Street, Second Floor
Palo Alto, CA 94306
(O) 650-424-0100 x 8413 (F) 650-424-0141